Performance Evaluation of Superstate HMM with Median Filter For Appliance Energy Disaggregation

Erwin Nashrullah
Department of Electrical Engineering
Faculty of Engineering
Universitas Indonesia
Depok, West Java
erwin.nashrullah@ui.ac.id

Abdul Halim
Department of Electrical Engineering
Faculty of Engineering
Universitas Indonesia
Depok, West Java
a.halim@ui.ac.id

***Abstract*— Information on electricity consumption is one of the essential elements in terms of regulating the distribution of electricity in smart micro grid. Besides, information on electricity consumption can help consumers carry out an evaluation process to reduce electricity bill costs, which indirectly affect overall energy efficiency. One method in the process of monitoring electricity consumption is Non-Intrusive Load Monitoring (NILM). The main problem in NILM is to determine the energy disaggregation consumed by several equipment by merely performing the retrieval of data from only one measuring point. We used the Superstate Hidden Markov Model as the tool for modelling and analysis. A median data filter to the input data is applied to improve the performance of the disaggregation process. Based on the results of tests conducted using the REDD, the lowest accuracy was 96.69% for all tests performed.**

Keywords*— Disaggregation, HMM, NILM, Median filter.

I. Introduction

Electric energy has become one of the basic needs of human life at this time. The increasing demand for electrical energy, which is not offset by the rise in the number of existing power plants requires a breakthrough in terms of energy development and the efficient use of electricity. The construction of a new large-scale power plant needs huge costs and more extended periods [1]. Taking into consideration that fossil fuel reserves will continuously tend to deplete and also the conventional power plant will produce carbon emissions into the atmosphere [2].

One solution to increasing electricity source is to build smart microgrid or community-based energy system. This system combines a community (or individually) owned electrical energy network with an information system that gives the status of energy availability to the primary energy network [3][4]. The use of smart meter will be an essential element for obtaining details of energy consumption in each electrical appliance. The consumed energy is recorded thru sensor, and all the data will be sent as parameters for analysis of electrical energy distribution. The recorded data will be also useful for consumers to take further action to use electrical energy appropriately to reduce the cost of the energy bill. Hence this will increase the efficiency of electrical energy distribution [5][6][7].

Two methods can be applied to provide data of electrical energy used in a house or a building. The first method is called intrusive load monitoring (ILM), and the other is Non-Intrusive Load Monitoring (NILM). On each electrical appliance, sensors are installed in the first method, while only one sensor is used in the NILM method for measuring the power of all connected appliances. The ILM requires much hardware and complex but straightforward algorithms. In contrast, NILM requires less hardware and lower installation cost, make researches on this field are exciting and growing fast [8]. Some successful method for disaggregation of power consumed by the cluster of electrical appliances are variations of HMM [9][10][11][12][13][14], variations of Neural Network [15][16][17], K-Nearest Neighbours (KNN) [18] and Support Vector Machine (SVM) [18]. Optimization techniques such as sparse matrix algorithm or filter application also improve performance of disaggregation.

A filter can be applied in the classification of electronic appliances to obtain good results. In this paper, we used the median filter of 5-window for reducing noise against our test data. The disaggregation method used was a variation of the Hidden Markov Model, namely Superstate Hidden Markov Model. We use data taken from REDD dataset. The data was divided into ten splits, and then ten-fold cross-validation was applied. We simulate the data in the noised mode, where the aggregate sum of the total energy of all appliances was not equal to the amount of power recorded on the power meter. We use the noised mode for the reason that this mode is the most realistic aggregation. Our test result shows that by applying a 5-window median filter, disaggregation with Super State Hidden Markov Model can increase the accuracy for the process of classification household appliances.

A filter can be applied in the classification of electronic appliances to obtain good results. In this paper, we used the median filter of 5-window for reducing noise against our test data [3]. The disaggregation method used was a variation of the Hidden Markov Model, namely Superstate Hidden Markov Model. We use data taken from REDD dataset [19]. The data was divided into ten splits, and then ten-fold cross-validation was applied. We simulate the data in the noised mode, where the aggregate sum of the total energy of all appliances was not equal to the amount of power recorded on the power meter. We use the noised mode for the reason that this mode is the most realistic aggregation. Our test result shows that by applying a 5-window median filter, disaggregation with Super State Hidden Markov Model can increase the accuracy for the process of classification household appliances.

II. Methodology

We divide our work into four consecutive steps, as shown in Fig 1. First, we applied 5-window median filter to our data in the preprocessing steps, and then we split data into two-part data for training and testing, respectively. Training data was used to build the SSHMM model.

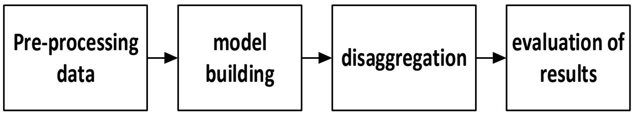


Fig. 1. Stages of the work process.

Based on this model, the disaggregation process is done for the noise mode with the help Viterbi algorithm. The last stage was for analysis of classification as the result of the applied filter.

A. Dataset

We used low-frequency data for our work. This data comes from REDD dataset[19]. The appliances used are refrigerators (REFG), dishwasher (appliance code: DISH), Kitchen Outlets (KTCH), lighting (LITE) taken from building data 1. The data used is 406.745 data for each electronic device.

B. Median Filter

The median filter is a non-linear mathematical tool to reduce noise contained in a time series data but still keeping its sharp, sustained change (edge). A median filter of 5-window is effective in minimizing the noise of the impulsive type. We use 5-window median filter because based on our previous work where we compare the performance of 3, 5, 7, and 9 window median filter. After comparing among them, we get result five window median filter produced a better performance to reduce noise compared to the other while maintaining the sharpness of the edge of signal data [3]. An algorithm of this kind of filter is shown in Fig. 2.

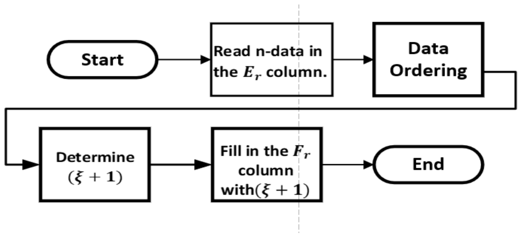


Fig. 2. Steps for applying the median n-window filter.

The algorithm starts by reading data size k in the column. E_r, r refers to the index of appliances whose the total number is M with 1 <r <M. Data from the E_r column will be processed as much as n data, which consists of data sets $\{y_1^{(1)} y_2^{(1)}, \; y_n^{(1)}\}$ then the data is arranged so that $y_1^{(1)} \leq y_2^{(1)}, \; \leq y_n^{(1)}$, the median value $\{y_1^{(1)} y_2^{(1)}, \; y_n^{(1)}\}$ is $(\xi + 1)$ order-t statistics with the value ξ adalah $\frac{(n-1)}{2}$ for m is odd. The median value obtained is stored in the F_r data column, and the process will be repeated until the entire median data from the E_r data column is obtained.

C. Determination of state for each appliance

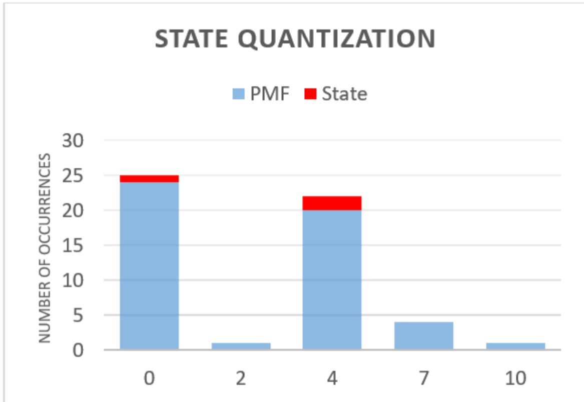
To build a superstate, the model requires a number of states present in each electronic appliance. Two steps are necessary to obtain equipment states, first from the observed output parameter we form its Probability Mass Function (PMF) and then secondly a quantization process is applied. PMF is a method often used to obtain discrete probability distribution for random scalar variable or multi-variant whose domain is discrete is discrete random variable in sample spaces, the PMF can be written as follows: : $X : S \quad A$ for $A \subseteq$ is a discrete random variable defined in the sample space S, then the probability of mass function $f(x) : A$ $[0, 1]$ for X can be defined Equation (1).

$$f(x) = \Pr(X = x) = P(\{s \in S : X(s) = x\}) \qquad (1)$$

Each value in the PMF is a positive value, and the total sum will produce a value of one $(\sum_{x \in A} f(x) = 1)$. The PMF was used in the preparation of this paper to describe the distribution of discrete distributions of observed variables in the form of electric current or the form of power from each electronic appliance. The value of consumption of electric current from each electronic device is obtained based on the dataset used. Equation (2) describes the PMF value of each appliance.

$$P_{mf} Y_m(n) = f(x) = \Pr[Y_m = n] n \in \{0, 1, 2 \ldots N_m\} \qquad (2)$$

$P_{mf} Y_m(n)$ is PMF for the m^{th} appliance at the observation value n with n is the observed observation value. After the PMF of each electronic appliance is obtained, the quantization process is performed on PMF data. This process divides the PMF data into several appliance work areas where the first region shows the electronic appliance part in the off condition and the next area shows the electronic appliance work area. Determining the state of work of electronic appliance is done by comparing the values of $P_{mf} Y_m(n)$ with $P_{mf} Y_m(n+1)$ and $P_{mf} Y_m(n \quad 1)$. If $P_{mf} Y_m(n \quad 1) < P_{mf} Y_m(n)$ and $P_{mf} Y_m(n) > P_{mf} Y_m(n+1)$ then that position is the peak of PMF and is a state from electronic appliance.



Observable value (N)	**0**	**2**	**4**	**7**	**10**
Count(n)	24	1	20	10	5
State Index	1	2			
Peak	0	4			
$\Pr(k^{(m)})$	0.4	0.6			

Fig. 3. PMF and quantization of electronic appliance.

Fig. 3 shows the PMF of 60 discrete data which produces 5 PMF values, but after quantization only two are detected as states (indicated on the histogram with a red label on top), particularly the position of the appliance at the value 0 and PMF at value 4. Position of the superstate (k_s) at time t can be known using equation (3).

$$k_s = k^{(M)} + \sum_{m=1}^{M-1} \left(k^{(m)} . \prod_{i=m+1}^{M} K^{(i)}\right) \qquad (3)$$

With $k^{(m)}$ is the internal quantization of the superstate for the m^{th} appliance. After the first two superstate conditions are known, the process of forming an SHMMF element can be carried out.

D. Building a model

In this paper, the process of creating electronic appliances models uses the superstate hidden Markov model method with the median filter (SSHMMF). SSHMMF is a hidden Markov model superstate (SSHMM) that has all the respect described in [12] combined using the median 5-window filter. The superstate is a combination of possible states that occur based on all the states that exist on each appliance. Fig. 3 shows an illustration of SSHMMF.

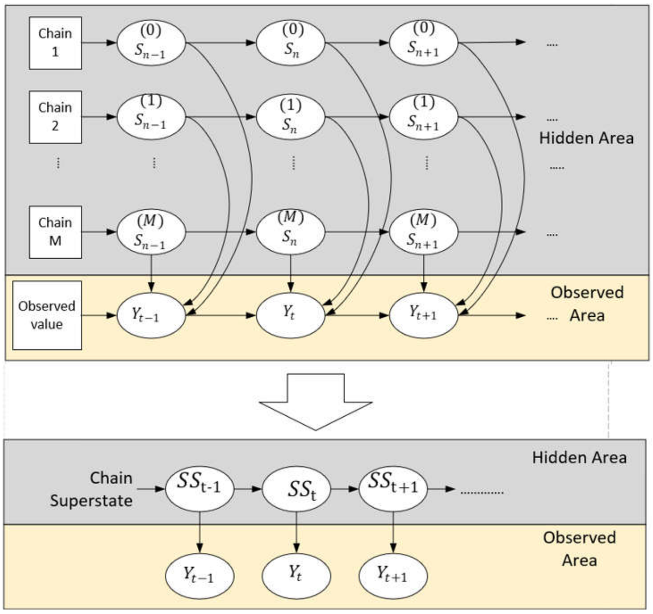


Fig. 4. Illustration of a Superstate Hidden Markov Model with a median filter.

As shown in Fig. 4, if there are M devices which each has a markov chain starting from the first chain to the M chain where M ∈ {1,2,3…M}and the number of observed data are N, with N= {1, 2, 3, …, N}. S_n^M is the hidden state of the M^{th}appl iancewhen the observation data is N. Observable parameters are denoted as Y_t, where Y_t is the parameter value observed at time t for observational data N. We define the Hidden Markov Superstate as the standard Hidden Markov Model [20], respectively $\lambda = (A,B,P_0)$, where A is the transition matrix between states, B is the observation matrix and P_0 is the initial probability matrix. If each electric appliance has the same state number K, then the superstate of all electronic appliances is G where G is $\prod_1^M K^{(m)}$ so that the matrix dimension A is GxG , matrix B is GxN and the initial probability is a vector with dimension $Gx1$. For example, if M is the number of the electronic appliances with M = {1,2,3, …M}, and N observation data in the form of total measured power for all electric appliances with N = $\{O_0, O_1, O_2, \ldots, O_N\}$. S_t is superstate at time t where $S_t = \{X_t^{(1)} X_t^{(2)} X_t^{(3)}, \ldots, X_t^{(M)}\}$ is given from the states of all M appliances. $X_t^{(m)}$ is internal state of the m^{th} appliance at time t. The probability of transition from present superstate (S_t) at time t to the next superstate (S_{t+1}) that finished at time t +1 is a_{ij} (i, j = 1,2,3,…G) that can be described using equation (4-6).

$$PA_{ij} = \begin{matrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1G} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2G} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3G} \\ \cdots & \cdots & \cdots & \cdots & \\ a_{G1} & a_{G2} & a_{G3} & \cdots & a_{GG} \end{matrix} \quad (4)$$

$$PB(O) = \begin{matrix} b_1(O_1) & b_1(O_2) & b_1(O_3) & b_1(O_4) & \cdots & b_1(O_N) \\ b_2(O_1) & b_2(O_2) & b_2(O_3) & b_2(O_4) & \cdots & b_2(O_N) \\ b_3(O_1) & b_3(O_2) & b_3(O_3) & b_3(O_4) & \cdots & b_3(O_N) \\ b_4(O_1) & b_4(O_2) & b_4(O_3) & b_4(O_4) & \cdots & b_4(O_N) \\ \cdots & \cdots & \cdots & \cdots & & \\ b_G(O_1) & b_G(O_2) & b_G(O_3) & b_G(O_4) & \cdots & b_G(O_N) \end{matrix} \quad (5)$$

$$P_0 = \begin{matrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \\ \\ \pi_G \end{matrix} \quad (6)$$

$b_i(O_j)$ is the probability of the occurrence of the O_j observation value at the time $t+1$ with the superstate condition is b_i where i={1, 2, 3,…G} and j = {1,2,3,…N}. Because observation value only depends on a particular superstate, then the value of $b_i(O_j) = 1$ for $i = j$ and zero otherwise. N is the observation value with N = $\{O_0, O_1, O_2, \ldots, O_N\}$. P_0 is a vector of the initial prior probabilities.

E. Sparse Viterbi

The sparse Viterbi algorithm is used to determine the state conditions that occur when the total electric current data at time *t* is given based on the SHMMF model that has been formed. This algorithm only calculates the non-zero element found in the matrix composing SHHMMF. The work process of the sparse Viterbi algorithm is almost the same as the general Viterbi algorithm shown in Fig. 5.

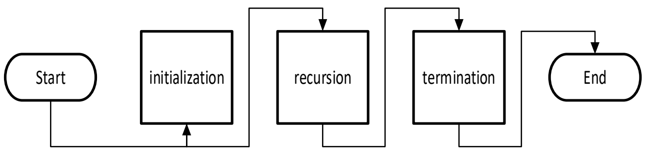


Fig.5. Stages of the Sparse Viterbi process.

The process begins with reading the data y_{t-1} and y_t where y_{t-1} is the total measured electric current of all electronic appliances at time *t-1*, and y_t is the measured total electric current data at time *t*. After obtaining y_{t-1} and y_t, the initialization process is performed by calculating the posterior probability for the first measured total electric power data (y_{t-1}). Posterior probability calculations are carried out using equation (7).

$$\delta_1(i) = P_0 b_i(O_t), 1 \leq i \leq K \quad (7)$$

The recursion process is carried out by performing posterior probability calculations at time *t* using equation (8).

$$\delta_t(j) = \max_{1 \leq i \leq N} [\delta_{t-1}(i)\, a_{ij}]\, b_j(O_t), 1 \leq j \leq K, 2 \leq t \leq T \quad (8)$$

The termination process is done by calculating the maximum value of the probability of the state occurring calculated based on equation (9).

$$P = \max_{1<j<K} [\delta_t(j)], I = \text{argmax}\, [\delta_t(j)] \quad (9)$$

The termination process will produce a value that indicates the highest possible superstate index that occurs when data is given.

III. Simulation and Analysis

Data processing for the simulation process is done in two ways based on processing of the noise in the data, namely the first mode is testing data without noise (denoised) and the second mode is testing data that contains noise (noised). Denoised is a test mode in which the total value of the measured power is equal to the sum of the power of each household electronic appliances while noised test mode is a test mode in which the noise in the data remained not removed. The amount of noise in the data used is the difference in the total power measured in total from the total power of all appliances. The data used are electronic appliances data in the form of a refrigerator (REFG), dishwasher (DISH), Kitchen Outlets (KTCH), lighting (LITE) obtained from REDD data in building 1. The amount of data used is 406,745 data divided into ten parts using 10-fold cross-validation so that for each test, ten models were obtained with different training and testing data to be analyzed.

Table I. Evaluation of model development.

Evaluation Aspect	Denoised		Noised	
	SSHMM	SSHMMF	SSHMM	SSHMMF
Train time	71.88 minutes	88.6 minutes	51.25 minutes	95.2 minutes
Train time/fold	7.19 minutes	8.86 minutes	5.12 minutes	9.52 minutes
Average Superstate	2931.2	2803.2	2931.2	2758.4
Size P_0	2931.2	2803.2	2931.2	2758.4
non-zero P_0	582.4	457.8	582.4	438.7
sparsity P_0	80,13%	83,66%	80,13%	84,09%
Storage requirement P_0	9606.4 bytes	7612.8 bytes	9606.4 bytes	7307.2 bytes
Size P_A	9006284.8	8293580.8	9006284.8	8051507.2
Non-zero P_A	3211	2204.4	3211	2156.4
sparsity P_A	99,96%	99,97%	0.99964347 1	99,97%
Storage requirement P_A	1756476.8 bytes	1341995.2 bytes	1756476.8 bytes	1258104.8 bytes
Size P_B	87938931. 2	84098803. 2	87938931.2	82754758. 4
Non-zero P_B	12257.8	9940.9	22182	21632.4
Sparsity P_B	99,98%	99,99%	99,97%	99,97%
Storage requirement P_B	4958102.4 bytes	4389468.8 bytes	13156976 bytes	12352444. 8 bytes

A. Model Analysis

Testing is done by forming a hidden Markov Model superstates with median filters and hidden Markov superstate models without a median filter in denoised and noised mode. Then we analyze the effect of adding a median filter to the aspect of the number of superstates, the size of the matrix, the number of non-zero values, the memory required to store the matrix P_A, P_B, P_0 and the length of time the modeling process is needed. Table I. shows the test results of the modeling SSHMM and SSHMMF.

Based on data Table I, It is known that the application of the median filter has the effect of decreasing the number of superstates produced. The reduced number of superstates will reduce the amount of non-zero values in the matrix P_A, matrix P_B, and vector P_0. The storage memory needed after adding the filter also becomes smaller.

B. Testing of disaggregation of electronic appliances

Testing is done by reading the total aggregate data from the measured active power to calculate the possible superstates that occur based on the models that have been made using the Sparse Viterbi algorithm. The disaggregation process is carried out after two aggregate data are read. The results of the disaggregation testing process for electronic appliances are shown in Table II.

Table II. Data classification result.

Test type	Denoised		Noised	
Parameter evaluated	SSHMM	SSHMMF	SSHMM	SSHMMF
Correct	158379	158223	156865	157294
Incorrect	4309	4465	5824	5393
TP	122397	122481	9724.5982	9446.304
TN	35982	35742	121462	121071
FP	1001	1795	35403	36223
FN	3308	2670	1580	1779
ATP	95991	96155	4244	3614
ITP	26405	26324	90721	91912
Inacc	8463.2804	8546.0145	30738	29157

For the validation process with the aim of measuring how accurately the estimation results are produced. Based on the data shown in Table II, we can calculate accuracy, F-score and Modified F-Score (M-F-score). The accuracy equation used is shown in equation (10).

$$Accuracy = \frac{correct}{correct+incorrect} \tag{10}$$

The equation for the F-score parameter is shown in equation (11)-(13).

$$Fscore = \frac{2p\ recision recall}{precision+recall} \tag{11}$$

$$precision = \frac{tp}{tp+fp} \tag{12}$$

$$recall = \frac{tp}{tp+fn} \tag{13}$$

Where precision is a positive predictive value and recall is a true positive rate or sensitivity, tp is true-positives (correctly predicted that the appliance is ON), fp is false-positive (appliance predicted to ON but OFF), and fn is false negative (the appliance is ON but is estimated to OFF). Other measurements of accuracy are introduced in [9], namely Modified Fcore (M-FScore) shown in equation (14)

$$MF_{score} = \frac{2M\ precision M\ recall}{M_{precision}+M_{recall}} \tag{14}$$

$$M_{precision} = \frac{ATP}{ATP+ITP+FP} \quad (15)$$

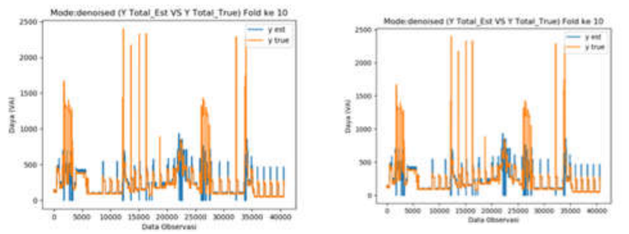
$$M_{recall} = \frac{ATP}{ATP+ITP+FN} \quad (16)$$

In M-Fscore measurement, True positive is divided into two parts, which are accurate true positive (ATP), and inaccurate true positive (ITP). The evaluation data result for the disaggregation process of SSHMM and SHMMF for denoised and noised mode are shown in Table III.

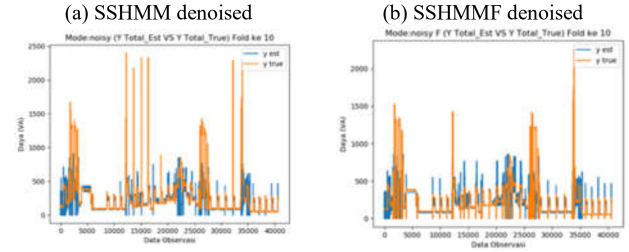
Table III. Evaluation data result for disaggregation test.

Evaluation parameter	Denoised		Noised	
	SSHMM	SSHMMF	SSHMM	SSHMMF
Accuracy	97.35%	97.26%%	96.42%	96.69%
Precision	99.19%	98.56%	98.72%	98.55%
Recall	97.37%	97.87%	96.62%	97.10%
F-Score	98.27%	98.21%	97.66%	97.82%
M Precision	77.79%	77.37%	73.73%	74.82%
M Recall	76.36%	76.83%	72.17%	73.72%
M F-Score	77.07%	77.10%	72.94%	74.27%

Based on the data shown in Table III, it can be seen that the addition of the median filter increases accuracy and M-Fscore in the noised test while the denoised test does not give too much influence. Accuracy shows the system's ability to recognize the condition of electronic equipment on or off, whereas M-F-score can be said to demonstrate the ability to predict electrical power from electronic equipment. The graph of the comparison of the results of the estimated total electric power with actual total electric power in the testing of electronic equipment disaggregation depicted in Fig.6.



(a) SSHMM denoised (b) SSHMMF denoised



(c) SSHMM noised (d) SSHMMF noised

Fig.6. Graph comparison between the prediction of electric power (estimation) and actual output (ground truth).

Fig. 6(a) and Fig. 6(b) show a comparison between the estimation of total electrical power for all electronic appliances (displayed using a blue line) with the ground truth (presented using orange lines) for the SSHMM and SSHMMF models in mode the data used is denoised mode. Fig. 6(c) and Fig. 6(d) show a comparison between the estimation of the total electrical power for all appliances with the ground truth for the SSHMM and SSHMMF models with the noised data mode. Based on Fig. 6, it can be seen that the median of the implanted filter can reduce noise in the signal without removing the signal edge characteristics. There is still a difference between the estimated value for electric power, and the measured value of electrical power shows the need for improvement in model optimization.

IV. Conclusion

Based on the results of the model tests that have been done, it can be concluded that the addition of the median filter on the hidden Markov model superstate has the effect of reducing the number of superstates and non-zero values, so that the PA matrix, PB and P0 vectors increase and affect the memory required for matrix storage. P_A, P_B and P_0 vectors are smaller than those with unfiltered superstates, but the formation of the Superstate hidden Markov model takes longer. The addition of the median filter also increases accuracy and M-FScore in the noised test mode. The Hidden Markov Superstate algorithm with the median filter successfully disaggregates a load of electronic equipment with the lowest accuracy of 96.69% and produces the lowest MF-Score of 72.98% of all test processes with variations in data models (denoised and noised tested with a REDD dataset.

Acknowledgment

This research is supported by Insentif Riset Sistem Inovasi Nasional (INSINAS) Tahun 2019 funded by Kementerian Ristek Diktik.

References

[1] T. Zia, D. Bruckner, and A. Zaidi, “A hidden Markov model based procedure for identifying household electric loads,” in *IECON Proceedings (Industrial Electronics Conference)*, 2011.

[2] K. A. Agyeman, S. Han, and S. Han, “Real-time recognition non-intrusive electrical appliance monitoring algorithm for a residential building energy management system,” *Energies*, 2015.

[3] E. Nashrullah and A. Halim, “Polynomial Load Model Development for Analysing Residential Electric Energy Use Behaviour,” *MATEC Web Conf.*, 2018.

[4] K. Basu, V. Debusschere, A. Douzal-Chouakria, and S. Bacha, “Time series distance-based methods for non-intrusive load monitoring in residential buildings,” *Energy Build.*, 2015.

[5] C. Fischer, “Feedback on household electricity consumption: A tool for saving energy?,” *Energy Effic.*, 2008.

[6] S. Darby, “The effectiveness of feedback on energy consumption,” 2006.

[7] R. Bonfigli, E. Principi, M. Fagiani, M. Severini, S. Squartini, and F. Piazza, “Non-intrusive load monitoring by using active and reactive power in additive Factorial Hidden Markov Models,” *Appl. Energy*, 2017.

[8] G. W. Hart, “Nonintrusive Appliance Load Monitoring,” *Proc. IEEE*, 1992.

[9] H. Kim, M. Marwah, M. Arlitt, G. Lyon, and J. Han, “Unsupervised Disaggregation of Low Frequency Power Measurements,” in *Proceedings of the 2011 SIAM International Conference on Data Mining*, 2011.

[10] A. Zoha, A. Gluhak, M. Nati, and M. A. Imran, “Low-power appliance monitoring using Factorial Hidden Markov Models,” in *Proceedings of the 2013 IEEE 8th International Conference on Intelligent Sensors, Sensor Networks and Information Processing: Sensing the Future, ISSNIP 2013*, 2013.

[11] O. Parson, S. Ghosh, M. Weal, and A. Rogers, “An unsupervised training method for non-intrusive appliance load monitoring,” *Artif. Intell.*, 2014.

[12] J. Z. Kolter and T. Jaakkola, “Approximate Inference in Additive Factorial HMMs with Application to Energy Disaggregation,” *Adv. Text. Technol.*, 2006.

[13] S. Makonin, F. Popowich, I. V. Bajic, B. Gill, and L. Bartram, “Exploiting HMM Sparsity to Perform Online Real-Time Nonintrusive Load Monitoring,” *IEEE Trans. Smart Grid*, 2015.

[14] I. Valera, F. Ruiz, L. Svensson, and F. Perez-Cruz, “Infinite Factorial Dynamical Model,” in *Advances in Neural Information Processing Systems 28*, 2015.

[15] J. Kelly and Knottenbelt W, “deep neural networks applied to energy disaggregation,” *Proc. 2nd ACM Int. Conf. Embed. Syst.*

energy-efficient built Environ. BuildSys '15. New York (NY, USA) ACM, pp. 55–64., 2015.

[16] L. Mauch and B. Yang, "A new approach for supervised power disaggregation by using a deep recurrent LSTM network," in *2015 IEEE Global Conference on Signal and Information Processing, GlobalSIP 2015*, 2016.

[17] L. Mauch and B. Yang, "A novel DNN-HMM-based approach for extracting single loads from aggregate power signals," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 2016.

[18] M. Figueiredo, A. de Almeida, and B. Ribeiro, "Home electrical signal disaggregation for non-intrusive load monitoring (NILM) systems," *Neurocomputing*, 2012.

[19] J. Z. Kolter and M. J. Johnson, "REDD : A Public Data Set for Energy Disaggregation Research," in *of the SustKDD Workshop on Data Mining Applications in Sustainability*, 2011.

[20] L. R. Rabiner and B. H. Juang, "An Introduction to Hidden Markov Models," *IEEE ASSP Mag.*, 1986.